Constructing Bayesian Network Graphs from Labeled Arguments

Remi Wieten[1(✉)], Floris Bex[1,2], Henry Prakken[1,3], and Silja Renooij[1]

[1] Information and Computing Sciences, Utrecht University, Utrecht, The Netherlands
{g.m.wieten,f.j.bex,h.prakken,s.renooij}@uu.nl
[2] Institute for Law, Technology and Society, Tilburg University, Tilburg, The Netherlands
[3] Faculty of Law, University of Groningen, Groningen, The Netherlands

Abstract. Bayesian networks (BNs) are powerful tools that are well-suited for reasoning about the uncertain consequences that can be inferred from evidence. Domain experts, however, typically do not have the expertise to construct BNs and instead resort to using other tools such as argument diagrams and mind maps. Recently, we proposed a structured approach to construct a BN graph from arguments annotated with causality information. As argumentative inferences may not be causal, we generalize this approach to include other types of inferences in this paper. Moreover, we prove a number of formal properties of the generalized approach and identify assumptions under which the construction of an initial BN graph can be fully automated.

Keywords: Bayesian networks · Argumentation · Inference · Reasoning

1 Introduction

Bayesian networks (BNs) [11] are compact graphical models of joint probability distributions that have found applications in many different fields where uncertainty plays a role, including medicine, forensics and law [6]. BNs are well-suited for reasoning about the uncertain consequences that can be inferred from evidence. However, especially in data-poor domains, their construction needs to be done mostly manually, which is a difficult, time-consuming and error-prone process [7], and domain experts typically resort to using other tools such as argument diagrams, mind maps and ontologies [4,8]. Hence, we believe BN construction can be facilitated by automatically extracting information relevant for a BN from such tools. More specifically, in this paper we study how information expressed as structured arguments [2] about the domain can inform the design of a BN *graph*, a directed acyclic graph (DAG) which captures the independence relation among variables.

In previous research, Bex and Renooij [3] identified constraints on BNs given structured arguments, but these only suffice for constructing an undirected skeleton of a BN graph. Recently, we were able to derive a directed graph [18], but only

G. Kern-Isberner and Z. Ognjanović (Eds.): ECSQARU 2019, LNAI 11726, pp. 99–110, 2019.
https://doi.org/10.1007/978-3-030-29765-7_9

by assuming that all inferences in the initial structured arguments are explicitly labeled with causality information [1,14]. Arcs in the BN graph are then set in the causal direction, following the heuristic typically used in the manual construction of BN graphs [11]. However, in [18] it is assumed that *all* inferences are labeled with causality information, which precludes the use of other types of inferences, such as mere statistical correlations and definitions. Furthermore, formal properties of the proposed proposals were not studied in [3,18].

Accordingly, in this paper we present an approach that generalizes our previously proposed construction approach [18] to other types of inference. In addition, we formally prove that BN graphs constructed by our approach allow reasoning patterns similar to the inferences represented in the original structured arguments. Moreover, we identify assumptions under which the fully automatically constructed initial graph is guaranteed to be a DAG, and we identify bounds on the complexity of inference in BNs constructed by our approach.

The paper is structured as follows. Section 2 provides preliminaries on argumentation and BNs. In Sect. 3, we present our generalized approach for constructing BN graphs from inferences. In Sect. 4, we prove a number of formal properties of the approach. In Sect. 5, we discuss related research and conclude.

2 Preliminaries

2.1 Argumentation

Throughout this paper, we assume that the domain experts' analysis is captured in an *argument graph* (AG), in which claims are substantiated by chaining inferences from the observed evidence; an example is depicted in Fig. 1a. AGs are closely related to argument diagrams and mind maps [4], familiar to many domain experts. Formally, an AG is a directed graph $G_{\mathcal{A}} = (\mathbf{P}, \mathbf{A}_{\mathcal{A}})$, where $\mathbf{P}$ is a set of nodes representing propositions from a literal language with ordinary negation symbol $\neg$, and $\mathbf{A}_{\mathcal{A}}$ is a set of directed (hyper)arcs. We write $p = -q$ in case $p = \neg q$ or $q = \neg p$. Nodes $\mathbf{E_p} \subseteq \mathbf{P}$ corresponding to the (observed) *evidence* are root nodes in $G_{\mathcal{A}}$. We assume that for every $p \in \mathbf{E_p}$ it holds that $\neg p \notin \mathbf{P}$. $\mathbf{A}_{\mathcal{A}}$ is comprised of three pairwise disjoint sets $\mathbf{S}$, $\mathbf{R}$ and $\mathbf{U}$, which are sets of support arcs, rebuttal arcs and undercutter arcs, respectively. A *support arc* is a (hyper)arc $s\colon \{p_1, \ldots, p_n\} \rightarrow p \in \mathbf{S}$, indicating an inference step from $\{p_1, \ldots, p_n\} \subseteq \mathbf{P}$ (called the *tails* of s, denoted by $\mathbf{Tails}(s)$) to a single proposition $p \in \mathbf{P}$ (called the *head* of s, denoted by $head(s)$). Here, curly brackets are omitted in case $|\mathbf{Tails}(s)| = 1$. Support arcs $s_1, \ldots, s_m$ form a *support chain* $(s_1, \ldots, s_m)$ iff $head(s_i) \in \mathbf{Tails}(s_{i+1})$ for $1 \leq i < m$.

There are two types of attack arcs. A *rebuttal arc* $r \in \mathbf{R}$ is a bidirectional arc $r\colon p \longleftrightarrow \neg p$ in $G_{\mathcal{A}}$ that exists for every pair $p, \neg p \in \mathbf{P}$. An *undercutter arc* $u \in \mathbf{U}$ is a hyperarc $u\colon p \rightarrow (s)$, where $p \in \mathbf{P}$ undercuts $s \in \mathbf{S}$. Informally, a rebuttal is an attack on a proposition, while an undercutter attacks an inference by providing exceptional circumstances under which the inference may not be applicable. In figures in this paper, nodes in $G_{\mathcal{A}}$ corresponding to elements of

$\mathbf{E_p}$ are shaded. Support arcs are denoted by solid (hyper)arcs and rebuttal arcs and undercutter arcs are denoted by dashed (hyper)arcs.

In reasoning about evidence, a distinction can be made between causal and evidential inferences [1,14]. Causal inferences are of the form "*c is a cause for e*" (e.g. fire causes smoke), whereas evidential inferences are of the form "*e is caused by c*" (e.g. smoke is caused by fire). Inferences may also be neither causal nor evidential. For instance, definitions, or abstractions [5], allow for reasoning at different levels of abstraction, such as stating that guns can generally be considered deadly weapons. Another example of a different type of inference is an inference representing a mere statistical correlation, such as a correlation between homelessness and criminality. While there may be one or more confounding factors that cause both homelessness and criminality (e.g. unemployment), a domain expert may be unaware of these factors or may wish to refrain from capturing them in the AG. For our current purposes, we assume that support arcs in $\mathbf{S}$ are either annotated with a causal "c" label, an evidential "e" label, or are labeled "o" for all other types of inferences. $\mathbf{S}$ then divides into three disjoint sets $\mathbf{S^c}$, $\mathbf{S^e}$ and $\mathbf{S^o}$ of causal, evidential and other types of support arcs, respectively. In figures in this paper, "o" labels are omitted.

In this paper, some further assumptions are made. We assume that support chains are *non-repetitive* in that there does not exist a support chain $(s, \ldots, s)$ in AG. We assume that for every support chain $(s_1, \ldots, s_n)$ the heads of $s_1, \ldots, s_n$ are *consistent* in that $\nexists i, j \in \{1, \ldots, n\}$, $i \neq j$ such that $head(s_i) = -head(s_j)$. Furthermore, we assume that AGs do not include *causal cycles* in that there do not exist two support chains $(s_1, \ldots, s_n)$ and $(s'_1, \ldots, s'_m)$ in AG with $s_1, \ldots, s_n \in \mathbf{S^c}$, $s'_1, \ldots, s'_m \in \mathbf{S^e}$, $\mathbf{Tails}(s_1) \cap \mathbf{Tails}(s'_1) \neq \emptyset$ and $head(s_n) = head(s'_m)$ or $head(s_n) = -head(s'_m)$. Informally, this assumption says that for every $p, q \in \mathbf{P}$, if p is a cause of q, then q (or $-q$) cannot be a cause of p (see also [1]).

As noted by Pearl [14], the chaining of a causal inference and an evidential inference can lead to undesirable results. Consider the example in which a causal inference states that a smoke machine causes smoke and an evidential inference states that smoke is evidence for fire. Chaining these inferences would make us conclude there is a fire when seeing a smoke machine, which is clearly undesirable. We therefore assume that an AG does not include a support chain (s_1, s_2) where $s_1 \in \mathbf{S^c}$, $s_2 \in \mathbf{S^e}$, and refer to this assumption as *Pearl's C-E constraint*.

For those familiar with argumentation, we note that, although we use the term "argument graph", the graph only represents inferences and attacks between propositions by means of arcs; actual arguments are not represented in the graph. Preferences over arguments, as well as their status, are thus not taken into account in our formalism, since they are not needed for our current purposes. Our formalism can be straightforwardly mapped to ASPIC$^+$ (cf. [2]) if all inferences are considered to be defeasible.

2.2 Bayesian Networks

A BN [11] compactly represents a joint probability distribution $\Pr(\mathbf{V})$ over a finite set of discrete random variables $\mathbf{V}$; in this paper we assume all variables to be Boolean. The variables are represented as nodes in a *DAG* $G_{\mathcal{B}} = (\mathbf{V}, \mathbf{A}_{\mathcal{B}})$, where $\mathbf{A}_{\mathcal{B}} \subseteq \mathbf{V} \times \mathbf{V}$ is a set of directed arcs $V_i \to V_j$ from parent V_i to child V_j. The BN further includes, for each node, a conditional probability table (CPT) specifying the probabilities of the values of the node conditioned on the possible joint value combinations of its parents. A node is called instantiated iff it is set to a specific value. Given a set of instantiations, or evidence, for nodes $\mathbf{E_V} \subseteq \mathbf{V}$, the probability distributions over the other nodes in the network can be updated through *probabilistic inference* [11]. An example of a BN graph is depicted in Fig. 1b, where ovals represent nodes and instantiated nodes are shaded.

The BN graph $G_{\mathcal{B}}$ captures the independence relation among its variables. Let a *chain* be defined as a sequence of distinct nodes and arcs in the BN graph. A node V is called a *head-to-head node* on a chain c if it has two incoming arcs on c. A chain c between nodes V_1 and V_2 is *blocked* iff it includes a node $V \notin \{V_1, V_2\}$ such that (1) V is an uninstantiated head-to-head node on c without instantiated descendants; or (2) V is instantiated and has at most one incoming arc on c. A chain that is not blocked is called *active*. If no active chains exist between V_1 and V_2 given instantiations of $\mathbf{Z} \subseteq \mathbf{V}$, then they are considered conditionally independent given $\mathbf{Z}$. In case a head-to-head node or one of its descendants is instantiated, an active chain is *induced* between its parents, allowing for interparental interactions. If one of the parents is now true, then the probability of another parent being true as well may change, depending on the specific synergistic effect modeled in the CPT for the head-to-head node.

BN construction is typically an iterative process. After constructing an initial BN graph, we should verify that this graph is acyclic and that it correctly captures the (conditional) independencies. If the graph does not yet exhibit these properties, arcs should be reversed, added or removed by the BN modeler in consultation with the domain expert. We call this the "*graph validation step*".

3 Constructing BN Graphs from Argument Graphs

To facilitate the BN construction process, we previously proposed a stepwise approach for constructing an initial BN graph from domain knowledge represented in AGs with support arcs in $\mathbf{S^c} \cup \mathbf{S^e}$ only [18]. In this section, we generalize this approach to include inferences in $\mathbf{S^o}$.

Upon using an AG to inform BN construction, we have to consider their difference in semantics. An AG, by means of its support chains, describes the iterative inference steps that can be made from the observed evidence towards the conclusions. In comparison, a BN describes a joint probability distribution which does not model such directionality. Only when probabilistic inference is performed is available evidence propagated through the network using the existing active chains. To mimic the inferences described by an AG in a BN, we will focus on ensuring that the (chains of) support arcs in the AG, originating from evidence $\mathbf{E_p} \subseteq \mathbf{P}$, are captured in the BN graph by means of active chains for propagating instantiations of $\mathbf{E_V} \subseteq \mathbf{V}$ (see also [18]). Note that since the notion of an active chain is a symmetrical concept, a BN graph will also capture reasoning patterns in the direction opposite of the support chains present in the

AG. In Sect. 4, we formally prove that all support chains in an AG indeed have corresponding active chains in the BN when following our generalized approach.

In the manual construction of BN graphs, arcs are typically directed using the *notion of causality* as a guiding principle [11]. By following this heuristic, two competing causes form a head-to-head connection in the node corresponding to the common effect, allowing synergistic effects between the causes to be directly captured in the CPT for this node. Hence, we propose to use the same heuristic in automatically directing arcs, where we exploit causality information explicitly expressed in an AG by means of "c" and "e" labels.

Undercutters attack inferences in support chains by providing exceptions to the inference. For instance, if an inference is in the evidential direction, then an undercutter suggests an alternative cause for the same effect. Accordingly, we propose to enable capturing such interactions between an undercutter and a support arc in the CPT of a head-to-head node formed in the BN graph.

3.1 The Generalized Approach

In this subsection, we present and explain the steps of the generalized approach. Let $var\colon \mathbf{P} \rightarrow \mathbf{V}$ be an operator mapping every proposition p or $\neg p \in \mathbf{P}$ in an AG to a BN variable $var(p) = var(\neg p) \in \mathbf{V}$ describing values p and $\neg p$. For an AG $G_{\mathcal{A}} = (\mathbf{P}, \mathbf{A}_{\mathcal{A}})$, a BN graph $G_{\mathcal{B}} = (\mathbf{V}, \mathbf{A}_{\mathcal{B}})$ is constructed as follows:

(1) $\forall p, \neg p \in \mathbf{P}$, include $var(p)$ in $\mathbf{V}$; if p or $\neg p \in \mathbf{E_p}$, also include $var(p)$ in $\mathbf{E_V}$.
(2) For every support arc $s\colon \{p_1, \ldots, p_n\} \rightarrow p$:
  (2a) If $s \in \mathbf{S^e}$, include $var(p) \rightarrow var(p_i)$, $i = 1, \ldots, n$ in $\mathbf{A}_{\mathcal{B}}$.
  (2b) If $s \in \mathbf{S^c}$, include $var(p_i) \rightarrow var(p)$, $i = 1, \ldots, n$ in $\mathbf{A}_{\mathcal{B}}$.
  (2c) If $s \in \mathbf{S^o}$ and $\nexists s_1 \in \mathbf{S^e}$ such that (s, s_1) form a support chain, include $var(p_i) \rightarrow var(p)$, $i = 1, \ldots, n$ in $\mathbf{A}_{\mathcal{B}}$.
  (2d) If $s \in \mathbf{S^o}$ and $\exists s_1, \ldots, s_m \in \mathbf{S^e}$ such that $(s_1, \ldots, s_m)$ is a maximal chain of evidential support arcs in AG following s, include $var(p_i) \rightarrow var(head(s_m))$, $i = 1, \ldots, n$ in $\mathbf{A}_{\mathcal{B}}$.
(3) For every undercutter arc $u\colon p \rightarrow (s) \in \mathbf{U}$ with $s\colon \{q_1, \ldots, q_n\} \rightarrow q$:
  (3a) If $s \in \mathbf{S^e}$, include $var(p) \rightarrow var(q_i)$, $i = 1, \ldots, n$ in $\mathbf{A}_{\mathcal{B}}$.
  (3b) If $s \notin \mathbf{S^e}$, include $var(p) \rightarrow var(q)$ in $\mathbf{A}_{\mathcal{B}}$.
(4) Verify the properties of the constructed graph $G_{\mathcal{B}}$:
  (4a) Break cycles in $G_{\mathcal{B}}$ introduced by so-called *evidential shortcuts* resulting from the combination of steps $2a$ and $2d$ (see Sect. 3.3 for further details).
  (4b) Apply the standard graph validation step (see Sect. 2.2).

While our approach exploits the domain knowledge captured in the AG in constructing a BN graph, the AG may lack information needed to prevent cycles and unwarranted (in)dependencies in the obtained BN graph; hence the manual validation step (step $4b$ above), which is standard in BN construction.

The first step is to capture every proposition in $G_{\mathcal{A}}$ and its negation as two values of a random variable in $G_{\mathcal{B}}$. By the same step, two propositions involved in a rebuttal are captured as two mutually exclusive values of the same node.

The steps pertaining to $s \in \mathbf{S^c} \cup \mathbf{S^e}$ are analogous to those proposed previously in [18]. These steps formalize the approach of setting arcs using the notion of causality as a guiding principle [11]. For further details, the reader is referred to [18]. In Sects. 3.2 and 3.3, we motivate and explain the steps pertaining to $s \in \mathbf{S^o}$ with several examples.

3.2 Explanation and Motivation of Steps 2*c* and 3*b*

Consider Fig. 1, illustrating steps 1−2*c* and 3*b* of the generalized approach for a forensic example. A dead body was found and we are interested in the cause of death of this person. According to witness testimony (*tes1*), the person was hit with a hammer (*hammer*); however, according to another testimony (*tes2*), the person was hit with a stone (*stone*). We conclude that the person was hit with an angular object (*angular*), as hammers and stones can generally be considered to be angular. Note that the relation between *hammer* (*stone*) and *angular* is neither causal nor evidential; instead, the support arcs between these propositions express that, at a higher level of abstraction, both hammers and stones can generally be considered angular objects. A mallet was found at the crime scene (*mallet*), which undercuts inference *hammer* → *angular* since a mallet is an exceptional type of hammer that is not angular but instead has a large cylindrical head. Finally, an autopsy report (*autopsy*) further supports the claim that the person was hit with an angular object. By following steps 1−3 of the generalized approach, the BN graph of Fig. 1b is constructed from the AG in Fig. 1a. By steps 2*c* and 3*b*, variables Hammer and Stone and variables Mallet and Hammer respectively form head-to-head connections in Angular.

In general, by step 2*c* head-to-head nodes are formed in the nodes corresponding to the heads of support arcs in $\mathbf{S^o}$. Specifically, let $p_1, \ldots, p_n$ be tails of one or more $s_i \in \mathbf{S^o}$ with $head(s_i) = p$. Then $\mathbf{A}_\mathcal{B}$ includes arcs $var(p_j) \to var(p)$, $j = 1, \ldots, n$ by step 2*c*; head-to-head nodes are, therefore, formed in $var(p)$. By setting arcs as per step 2*c*, we thus allow for including synergistic effects, if any, of the tails on the probability of p in the CPT for the head-to-head node.

Similarly, by step 3*b* head-to-head nodes are formed in the nodes corresponding to the heads of undercut support arcs in $\mathbf{S^c} \cup \mathbf{S^o}$. Specifically, let $u\colon p \to (s) \in \mathbf{U}$ be an undercutter of $s\colon \{q_1, \ldots, q_n\} \to q \in \mathbf{S^c} \cup \mathbf{S^o}$. Then by step

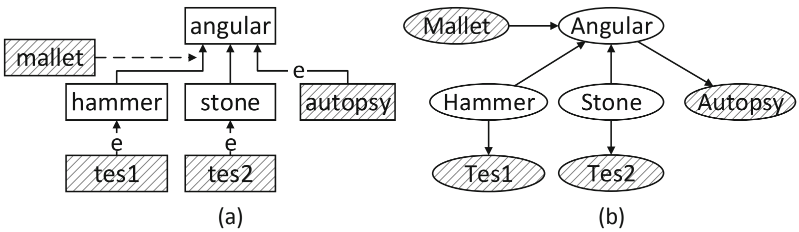

Fig. 1. An AG including support arcs in $\mathbf{S^o}$ (a); the corresponding BN graph constructed by steps 1−2*c* and 3*b* of the generalized approach (b).

$3b$, head-to-head nodes are formed in $var(q)$ as $\mathbf{A}_{\mathcal{B}}$ includes arc $var(p) \rightarrow var(q)$. Again, this allows for modeling possible interactions between p and q_i, and hence between $var(p)$ and $var(q_i)$, directly in the CPT for $var(q)$. Bex and Renooij [3] previously noted that the presence of an undercutter should decrease the probability that the conclusion of the undercut inference is true. By setting arcs as per step $3b$, this interaction can be directly captured by the following constraints on the CPT for $var(q)$: $\Pr(q \mid p, q_i) < \Pr(q \mid \neg p, q_i)$ for $i = 1, \ldots, n$.

3.3 Explanation and Motivation of Steps 2*d* and 4*a*

Next, consider Figs. 2a and b, illustrating step $2d$ of the generalized approach for a medical example (taken from [7]). After performing a CT scan (*scan*) on a patient who has severe difficulty swallowing, it is established that a tumor is present in the lower (*distal*) part of his esophagus. Clinical studies indicate a strong correlation between the location of an esophageal tumor and its cell type; however, neither can be considered a cause of the other. Distal tumors generally consist of cylindrical cells (*cylindrical*), often formed as a result of frequent gastric reflux (*reflux*). The BN graph constructed by steps 1–$2a$ and $2d$ of the generalized approach from the AG in Fig. 2a is depicted in Fig. 2b. As arcs Distal $\rightarrow$ Reflux and Reflux $\rightarrow$ Cylindrical are included in $\mathbf{A}_{\mathcal{B}}$ and the involved nodes are not instantiated, active chains exist between Distal and Reflux and Distal and Cylindrical. Note that we do not wish to set arcs as per step $2c$, as in this case a head-to-head node would instead be formed in Cylindrical which would block the chain between Distal and Reflux.

Under specific conditions, cycles are introduced in step $2d$ of the generalized approach, namely when a so-called *evidential shortcut* exists in the AG, i.e. if in addition to the conditions of step $2d$, also $\exists s'_1, \ldots, s'_k \in \mathbf{S^e}$ such that $(s'_1, \ldots, s'_k)$ form a support chain, $\mathbf{Tails}(s) \cap \mathbf{Tails}(s'_1) \neq \emptyset$ and $head(s'_k) = head(s_j)$ or $head(s'_k) = -head(s_j)$ for a $j \in \{1, \ldots, m\}$. An example is depicted in Fig. 2c. In this example, $s\colon p \rightarrow q_1 \in \mathbf{S^o}$ is followed by a chain of support arcs $s_1\colon q_1 \rightarrow r, s_2\colon r \rightarrow s \in \mathbf{S^e}$, where there also exists a chain of support arcs $s'_1\colon p \rightarrow$

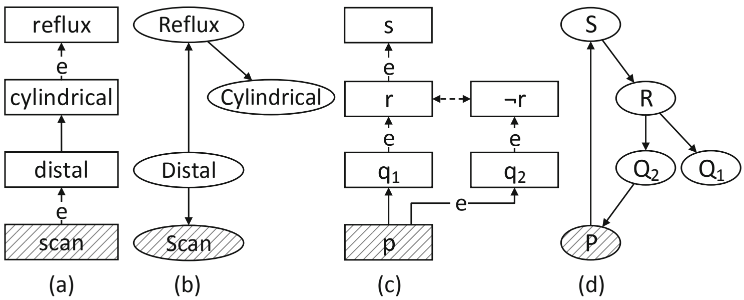

Fig. 2. An AG (a) and the corresponding BN graph (b) illustrating step $2d$ of the generalized approach; an AG (c) and the corresponding BN graph (d), illustrating the conditions under which a cycle is introduced in step $2d$.

q_2, s_2': $q_2 \rightarrow \neg r \in \mathbf{S^e}$. By step $2a$, arcs $S \rightarrow R, R \rightarrow Q_2$ and $Q_2 \rightarrow P$ are included in $\mathbf{A}_\mathcal{B}$. By step $2d$, arc $P \rightarrow S$ is also included, introducing a cycle in $G_\mathcal{B}$. We note that this arc can safely be removed, as an active chain already exists between P and S via Q_2 and R. In general, cycles are broken in step $4a$ by removing arcs $var(p_l) \rightarrow var(head(s_m))$ from $\mathbf{A}_\mathcal{B}$ $\forall p_l \in \mathbf{Tails}(s) \cap \mathbf{Tails}(s_1')$.

4 Properties of the Generalized Approach

In this section, we prove a number of formal properties of the generalized approach. The first property states that for every support chain in a given AG there indeed exists a corresponding active chain in the BN graph.

Proposition 1. *Let $G_\mathcal{A} = (\mathbf{P}, \mathbf{A}_\mathcal{A})$ be an AG with root nodes $\mathbf{E_p}$, and let $G_\mathcal{B} = (\mathbf{V}, \mathbf{A}_\mathcal{B})$ be the corresponding BN graph constructed according to steps 1–$4a$ of the generalized approach. Let $(s_1, \ldots, s_n)$ be any support chain in $G_\mathcal{A}$, where $\mathbf{Tails}(s_1) = \{p_1, \ldots, p_m\}$ and $head(s_n) = q$. Then there exist active chains between $var(p_i)$ and $var(q)$ in $G_\mathcal{B}$ given $\mathbf{E_V}$ for every $i \in \{1, \ldots, m\}$.*

Proof (Sketch). The following cases are distinguished:

- If $s_k \in \mathbf{S^c} \cup \mathbf{S^e}$ $\forall k \in \{1, \ldots, n\}$, then when following steps $2a$ and $2b$ a head-to-head node can only be formed in $var(head(s_j))$ for an arbitrary s_j, $j \in \{1, \ldots, n-1\}$ if $s_j \in \mathbf{S^c}$, $s_{j+1} \in \mathbf{S^e}$; however, this construction is prohibited as it violates Pearl's C-E constraint (see Sect. 2.1). Furthermore, since heads of support arcs are not propositions in $\mathbf{E_p}$, corresponding nodes in $G_\mathcal{B}$ are not instantiated. Chains between $var(p_i)$ and $var(q)$ are thus never blocked.
- If $(s_1, \ldots, s_n)$ includes support arcs in $\mathbf{S^o}$ and none of these arcs is followed by an $s \in \mathbf{S^e}$, then arcs in $\mathbf{A}_\mathcal{B}$ are set similarly as for $s \in \mathbf{S^c}$ by step $2c$. As per the above proof, chains are not blocked.
- Let an $s_j \in \mathbf{S^o}$, $1 \leq j < n$ be followed by a chain of support arcs in $\mathbf{S^e}$, and let $(s_{j+1}, \ldots, s_{j+l})$ be a maximal such chain. If $j + l \leq n$, then step $2d$ introduces direct arcs, and therefore active chains, between nodes in $\{var(p) \mid p \in \mathbf{Tails}(s_j)\}$ and $var(head(s_{j+l}))$. If $j + l > n$, then $\mathbf{A}_\mathcal{B}$ in addition includes a directed path from $var(head(s_{j+l}))$ to $var(head(s_n))$ by step $2a$; therefore, chains between nodes in $\{var(p) \mid p \in \mathbf{Tails}(s_j)\}$ and $var(head(s_n))$ via $var(head(s_{j+l}))$ are active, as $var(head(s_{j+l}))$ is not a head-to-head node. In step $4a$, a subset of the arcs introduced in step $2d$ is removed (see Sect. 3.3) iff an evidential shortcut and a corresponding active chain already exist.

Finally, $\mathbf{A}_\mathcal{B}$ is only extended for undercutter arcs in step 3; active chains formed between $var(p_i)$ and $var(q)$ in step 2 are, therefore, not affected by this step. $\square$

In Proposition 2, we prove that under specific conditions on AGs an acyclic graph is automatically obtained when following steps 1–$4a$ of the approach, which simplifies the manual verification involved in step $4b$. Conditions (a) and (b) concern the existence of undercutter arcs within and between connected subgraphs of AGs. Condition (c) is a generalization of our assumption that no causal cycles exist in AGs (see Sect. 2.1) to support arcs in $\mathbf{S^o}$.

Proposition 2. *Let $G_{\mathcal{A}} = (\mathbf{P}, \mathbf{A}_{\mathcal{A}})$, and let $G^*_{\mathcal{A}} = (\mathbf{P}, \mathbf{A}^*_{\mathcal{A}})$ be the subgraph of $G_{\mathcal{A}}$ with $\mathbf{A}^*_{\mathcal{A}} = \mathbf{A}_{\mathcal{A}} \setminus \mathbf{U}$. Let an* AG component of $G_{\mathcal{A}}$ *be defined as a connected component of $G^*_{\mathcal{A}}$. Assume the following conditions are satisfied:*

(a) For any AG component $C = (\mathbf{P}', \mathbf{A}'_{\mathcal{A}})$ of $G_{\mathcal{A}}$ with $\mathbf{P}' \subseteq \mathbf{P}$, $\mathbf{A}'_{\mathcal{A}} \subseteq \mathbf{A}^*_{\mathcal{A}}$, there does not exist a $u\colon p \to (s) \in \mathbf{U}$ with $p \in \mathbf{P}'$, $s \in \mathbf{A}'_{\mathcal{A}}$.

(b) For every pair of AG components $C_1 = (\mathbf{P}', \mathbf{A}'_{\mathcal{A}})$ and $C_2 = (\mathbf{P}'', \mathbf{A}''_{\mathcal{A}})$ of $G_{\mathcal{A}}$ with $\mathbf{P}', \mathbf{P}'' \subseteq \mathbf{P}$, $\mathbf{A}'_{\mathcal{A}}, \mathbf{A}''_{\mathcal{A}} \subseteq \mathbf{A}^*_{\mathcal{A}}$, there does not exist both a $u_1\colon p_1 \to (s_1) \in \mathbf{U}$ with $p_1 \in \mathbf{P}'$, $s_1 \in \mathbf{A}''_{\mathcal{A}}$ and a $u_2\colon p_2 \to (s_2) \in \mathbf{U}$ with $p_2 \in \mathbf{P}''$, $s_2 \in \mathbf{A}'_{\mathcal{A}}$.

(c) There do not exist two support chains $(s_1, \ldots, s_n)$ and $(s'_1, \ldots, s'_m)$ with $s_1, \ldots, s_n \in \mathbf{S^c} \cup \mathbf{S^o}$, $s'_1, \ldots, s'_m \in \mathbf{S^e}$, $\mathbf{Tails}(s_1) \cap \mathbf{Tails}(s'_1) \neq \emptyset$, and $head(s_n) = head(s'_m)$ or $head(s_n) = -head(s'_m)$.

Let $G_{\mathcal{B}} = (\mathbf{V}, \mathbf{A}_{\mathcal{B}})$ be the graph constructed from $G_{\mathcal{A}}$ according to steps 1*–*4*a of the generalized approach. Then $G_{\mathcal{B}}$ is a DAG.*

Proof (Sketch). The following cases are distinguished:

- In steps 2*a* and 2*b*, no cycles are introduced. Specifically, our non-repetitiveness assumption and our consistency assumption (see Sect. 2.1) jointly assume that for every $p \in \mathbf{P}$, p or $-p$ cannot be inferred via a chain of support arcs. Therefore, no chain of arcs exists in $\mathbf{A}_{\mathcal{B}}$ from a node P to itself. The only other case in which cycles can be introduced is when a causal cycle exists in $G_{\mathcal{A}}$, which is also prohibited by assumption (see Sect. 2.1).
- No cycles are introduced in step 2*c* if condition (c) is satisfied. Cycles are only introduced in step 2*d* if an evidential shortcut exists; however, these cycles are broken again in step 4*a* as described in Sect. 3.3.
- After step 2, there is a correspondence between AG components and the connected components of the underlying undirected graph S of the thus far constructed BN graph. Under condition (a), no cycles are introduced within a connected component of S when including additional arcs in $\mathbf{A}_{\mathcal{B}}$ for every $u \in \mathbf{U}$ in step 3. Furthermore, for every pair of AG components C_1 and C_2 of $G_{\mathcal{A}}$ with corresponding connected components C'_1 and C'_2 of S, no cycles are introduced between components C'_1 and C'_2 in step 3 under condition (b). □

Figures 3a and c depict examples of AGs that do not satisfy conditions (a) and (b) of Proposition 2, respectively. In the validation step that follows the initial construction of these BN graphs, arcs can be reversed or removed to make these graphs acyclic. The choice of arc to reverse or remove will depend on its effect on active chains, including those between nodes not directly incident on the arc. We note that this type of manual verification is standard in BN construction, especially in data-poor domains. While the domain knowledge expressed in the original AG has been exploited to construct an initial BN graph, additional domain knowledge may need to be elicited to obtain a valid graph.

Proposition 3 gives an upper-bound on the number of parents introduced by the approach for each node $var(p)$ in a BN graph, which bounds both the size of the CPTs and the complexity of inference in the BN. This bound captures the

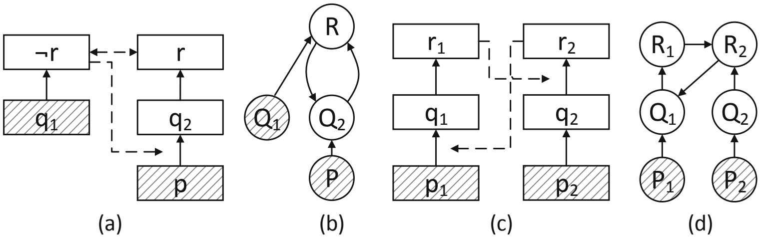

Fig. 3. Examples of AGs (a, c) for which a cyclic graph is constructed by steps 1−4a of the generalized approach (b, d).

number of support arcs and undercutters that involve either proposition p or $\neg p$. The proof of this result is straightforward and omitted due to space limitations.

Proposition 3. *Let $G_\mathcal{A} = (\mathbf{P}, \mathbf{A}_\mathcal{A})$ be an AG, and let $G_\mathcal{B} = (\mathbf{V}, \mathbf{A}_\mathcal{B})$ be the BN graph constructed according to steps* 1−4a *of the generalized approach. For every $p \in \mathbf{P}$, let $\mathbf{Par}_p = \{p_i \mid p_i \in \mathbf{Tails}(s),\ s \in \mathbf{S^c} \cup \mathbf{S^o},\ head(s) = p\}$ and let $\mathbf{Par}'_p = \{p_i \mid p_i \in \mathbf{Tails}(s),\ s \in \mathbf{S^o},\ s$ is followed by maximal chain $s_1, \ldots, s_m \in \mathbf{S^e}$ with $head(s_m) = p$ or $head(s_m) = \neg p\}$. Let $\mathbf{S}^\mathbf{e}_p$ be a subset of $\mathbf{S^e}$, where $s \in \mathbf{S}^\mathbf{e}_p$ iff $p \in \mathbf{Tails}(s)$. Let $\mathbf{U}^\mathbf{e}_p \subseteq \mathbf{U}$ be the subset of undercutter arcs directed to an $s \in \mathbf{S}^\mathbf{e}_p$ or $s \in \mathbf{S}^\mathbf{e}_{\neg p}$. Similarly, let $\mathbf{U}^\mathbf{c}_p, \mathbf{U}^\mathbf{o}_p \subseteq \mathbf{U}$ be the subsets of undercutter arcs directed to an $s \in \mathbf{S^c}$ respectively $\mathbf{S^o}$ for which $head(s) = p$ or $head(s) = \neg p$. Then an upper-bound for the number of parents of $var(p)$ is:*

(1) $|\mathbf{Par}_p| + |\mathbf{Par}_{\neg p}| + |\mathbf{Par}'_p| + |\mathbf{U}^\mathbf{c}_p| + |\mathbf{U}^\mathbf{o}_p|$ *if* $\mathbf{S}^\mathbf{e}_p = \mathbf{S}^\mathbf{e}_{\neg p} = \emptyset$;
(2) $|\mathbf{Par}_p| + |\mathbf{S}^\mathbf{e}_{\neg p}| + |\mathbf{Par}'_p| + |\mathbf{U}^\mathbf{e}_p| + |\mathbf{U}^\mathbf{c}_p| + |\mathbf{U}^\mathbf{o}_p|$ *if* $\mathbf{S}^\mathbf{e}_p = \emptyset$ *and* $\mathbf{S}^\mathbf{e}_{\neg p} \neq \emptyset$;
(3) $|\mathbf{Par}_{\neg p}| + |\mathbf{S}^\mathbf{e}_p| + |\mathbf{Par}'_p| + |\mathbf{U}^\mathbf{e}_p| + |\mathbf{U}^\mathbf{c}_p| + |\mathbf{U}^\mathbf{o}_p|$ *if* $\mathbf{S}^\mathbf{e}_p \neq \emptyset$ *and* $\mathbf{S}^\mathbf{e}_{\neg p} = \emptyset$;
(4) $|\mathbf{S}^\mathbf{e}_p| + |\mathbf{S}^\mathbf{e}_{\neg p}| + |\mathbf{U}^\mathbf{e}_p| + |\mathbf{U}^\mathbf{o}_p|$ *if* $\mathbf{S}^\mathbf{e}_p \neq \emptyset$ *and* $\mathbf{S}^\mathbf{e}_{\neg p} \neq \emptyset$.

5 Conclusion

In this paper, we have studied how domain knowledge expressed as labeled arguments can be exploited to construct a BN graph. Firstly, we have generalized our previously proposed approach [18] by allowing inference types that are neither causal nor evidential. Moreover, we have formally proven that, as intended, our approach captures all support chains in an AG in the form of active chains in the BN graph. We have also identified conditions on AGs under which a DAG is automatically constructed by the approach, simplifying the manual verification step. Lastly, we have identified bounds on the size of the CPTs and the complexity of inference in BNs constructed by our approach. All properties also hold for the limited case considered in [18] but were not proven in that paper.

The generalized approach allows us to construct an initial BN graph from a domain expert's initial argument-based analysis, capturing similar reasoning patterns as their original AG; it thereby simplifies the BN elicitation process.

We note that BN construction is an iterative process in which both the domain expert and BN modeler should stay involved; this also holds when applying our approach, as the provided AG may be incomplete or incorrect. To aid in this iterative process, approaches were proposed in related work which allow experts to use argumentation to argue about the BN under construction instead of about the domain [13,19]. In other related work, approaches for explaining the reasoning patterns captured in BNs in terms of argumentation were proposed [12,17], which allow domain experts more accustomed to argumentation to understand the probabilistic reasoning captured in a BN. Compared to the present paper, this work is in the reverse direction, namely from BNs to arguments.

Recently, there has been much other work on probabilistic argumentation. However, most approaches concern *abstract* argumentation (see e.g. [10] for an overview) while we need structured arguments. Rienstra [16] considers probabilistic structured argumentation; however, he takes what Hunter [9] calls the constellations approach to probabilistic argumentation by considering uncertainty in the existence of arguments. Instead, we take what Hunter calls the epistemic approach to probabilistic argumentation by considering probabilities to express uncertainty concerning the reliability of an argument's inferences. There is some work on the epistemic approach to probabilistic structured argumentation (e.g. [9,15]). In future work, this may become relevant for deriving probabilistic constraints on BNs.

References

1. Bex, F.: An integrated theory of causal stories and evidential arguments. In: Proceedings of the Fifteenth International Conference on Artificial Intelligence and Law, pp. 13–22. ACM Press, New York (2015)
2. Bex, F., Modgil, S., Prakken, H., Reed, C.A.: On logical specifications of the argument interchange format. J. Logic Comput. **23**(5), 951–989 (2013)
3. Bex, F., Renooij, S.: From arguments to constraints on a Bayesian network. In: Baroni, P., Gordon, T.F., Scheffler, T., Stede, M. (eds.) Computational Models of Argument: Proceedings of COMMA 2016, vol. 287, pp. 95–106. IOS Press, Amsterdam (2016)
4. Buckingham Shum, S.J.: The roots of computer supported argument visualization. In: Kirschner, P.A., Buckingham Shum, S.J., Carr, C.S. (eds.) Visualizing Argumentation: Software Tools for Collaborative and Educational Sense-Making, pp. 3–24. Springer, London (2003). https://doi.org/10.1007/978-1-4471-0037-9_1
5. Console, L., Dupré, D.T.: Abductive reasoning with abstraction axioms. In: Lakemeyer, G., Nebel, B. (eds.) Foundations of Knowledge Representation and Reasoning. LNCS (LNAI), vol. 810, pp. 98–112. Springer, Heidelberg (1994). https://doi.org/10.1007/3-540-58107-3_6
6. Fenton, N., Neil, M.: Risk Assessment and Decision Analysis with Bayesian Networks. CRC Press, Boca Raton (2012)
7. van der Gaag, L.C., Helsper, E.M.: Experiences with modelling issues in building probabilistic networks. In: Gómez-Pérez, A., Benjamins, V.R. (eds.) EKAW 2002. LNCS (LNAI), vol. 2473, pp. 21–26. Springer, Heidelberg (2002). https://doi.org/10.1007/3-540-45810-7_4

8. Helsper, E.M., van der Gaag, L.C.: Building Bayesian networks through ontologies. In: van Harmelen, F. (ed.) Proceedings of the Fifteenth European Conference on Artificial Intelligence, vol. 77, pp. 680–684. IOS Press, Amsterdam (2002)
9. Hunter, A.: A probabilistic approach to modelling uncertain logical arguments. Int. J. Approx. Reason. **54**(1), 47–81 (2013)
10. Hunter, A., Thimm, M.: On partial information and contradictions in probabilistic abstract argumentation. In: Baral, C., Delgrande, J., Wolter, F. (eds.) Principles of Knowledge Representation and Reasoning: Proceedings of the Fifteenth International Conference (KR-2016), pp. 53–62. AAAI Press, Palo Alto (2016)
11. Jensen, F.V., Nielsen, T.D.: Bayesian Networks and Decision Graphs, 2nd edn. Springer, New York (2007). https://doi.org/10.1007/978-0-387-68282-2
12. Keppens, J.: Argument diagram extraction from evidential Bayesian networks. Artif. Intell. Law **20**(2), 109–143 (2012)
13. Keppens, J.: On modelling non-probabilistic uncertainty in the likelihood ratio approach to evidential reasoning. Artif. Intell. Law **22**(3), 239–290 (2014)
14. Pearl, J.: Embracing causality in default reasoning. Artif. Intell. **35**(2), 259–271 (1988)
15. Prakken, H.: Probabilistic strength of arguments with structure. In: Thielscher, M., Toni, F., Wolter, F. (eds.) Principles of Knowledge Representation and Reasoning: Proceedings of the Sixteenth International Conference (KR-2018), pp. 158–167. AAAI Press, Palo Alto (2018)
16. Rienstra, T.: Towards a probabilistic Dung-style argumentation system. In: Ossowski, S., Toni, F., Vouros, G.A. (eds.) Proceedings of the First International Conference on Agreement Technologies, pp. 138–152. CEUR, Aachen (2012)
17. Timmer, S.T., Meyer, J.-J.C., Prakken, H., Renooij, S., Verheij, B.: Explaining Bayesian networks using argumentation. In: Destercke, S., Denoeux, T. (eds.) ECSQARU 2015. LNCS (LNAI), vol. 9161, pp. 83–92. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-20807-7_8
18. Wieten, R., Bex, F., Prakken, H., Renooij, S.: Exploiting causality in constructing Bayesian networks from legal arguments. In: Palmirani, M. (ed.) Legal Knowledge and Information Systems: JURIX 2018: The Thirty-First Annual Conference, vol. 313, pp. 151–160. IOS Press, Amsterdam (2018)
19. Wieten, R., Bex, F., Prakken, H., Renooij, S.: Supporting discussions about forensic Bayesian networks using argumentation. In: Proceedings of the Seventeenth International Conference on Artificial Intelligence and Law, pp. 143–152. ACM Press, New York (2019)